

ADA 076827

LEVEL *IV*

41

Research Memorandum 76-29

ACCEPTABILITY OF ASSOCIATE EVALUATION IN NATIONAL GUARD UNITS

Stephen L. Goldberg and Joseph F. Lombardo, Jr.

DDC
REFORM
NOV 19 1979
E

DDC FILE COPY



U. S. Army

Research Institute for the Behavioral and Social Sciences

December 1976

79 11 13 303

This document has been approved
for public release and sale; its
distribution is unlimited.

**Best
Available
Copy**

DISPOSITION FORM

For use of this form, see AR 340-15, the proponent agency is TAGCEN.

REFERENCE OR OFFICE SYMBOL	SUBJECT
PERI-TP	Clearance and Transmittal of Reports to DTIC

TO DDC-DAA-1
ATTN: Mr. Schrecengost

FROM ARI Rsch Pub Group
See the following

DATE 8 Nov 79
Ms Price/48913

CMT 1

1. The reports listed on Inclosure 1 are approved for public release with unlimited distribution (50 numbered ARI Research Memorandums, 74-1 thru 76-30).
2. These are among the previously unrecorded ARI reports which you identified to us 22 June 1979 as not in your retrieval system. The accompanying box contains at least one copy of each report for your retention and reproduction.

1 incl
List of reports, 1974-76

Helene S. Price

HELEN S. PRICE
Research Publications Group
Army Research Institute

Accession For	
THIS GROUP	<input checked="checked" type="checkbox"/>
DOO TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	<input type="checkbox"/>
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

Army Project Number
16 2Q763731A768

Performance-Based
Skill Development

9 *memo*
Research Memorandum 76-29

6 ACCEPTABILITY OF ASSOCIATE EVALUATION
IN NATIONAL GUARD UNITS.

10 Stephen L. Goldberg Joseph F. Lombardo, Jr.

Milton H. Maier, Work Unit Leader

Submitted by:
Milton S. Katz, Chief
INDIVIDUAL TRAINING AND SKILL EVALUATION TECHNICAL AREA

11 December 1976

Approved by:

14
ARI-RM-76-29

E. Ralph Dusek, Director
Individual Training and Performance
Research Laboratory

J. E. Uhlaner, Director
U.S. Army Research Institute for
The Behavioral and Social Sciences

Research Memorandums are informal reports on technical research problems. Limited distribution is made, primarily to personnel engaged in research for the Army Research Institute.

408 020

slt

ACCEPTABILITY OF ASSOCIATE EVALUATION IN NATIONAL GUARD UNITS

Associate ratings have been used by the U.S. Army as a measure of ability or leadership in a variety of training situations. Peer evaluations have been found to be reliable and valid predictors of performance in training contexts (Downey and Duffy, 1975). Reliability coefficients have consistently been in the range of .70 to .90, and the validity of the ratings has been found equal to or higher than that of other measures which have been used (Bolton, 1971; Downey, 1974, 1975; Downey, Medland, and Yates, 1976).

Given the nature of the training environments in which associate ratings have been used--class size units during most of the training time--and the use of computerized scoring techniques, peer ratings have been found to be feasible. They can be administered and scored with little disruption to routine. A possible drawback with the technique is that individuals who take part in the evaluations have in some situations found them to be an unacceptable form of assessment (Downey, 1975; Mohr, 1975). Downey (1976) found that Ranger students felt that they could evaluate the leadership potential of their classmates but preferred not to have peer ratings included in their records. A large sample of students in Officer Basic Courses (OBC) at various service schools reported that they felt that their experiences and observations were an inadequate basis for judging others' ability. Their attitudes about use of the ratings were less negative, but they still felt it would be unacceptable to use ratings for career decisions (Mohr, 1975).

All the research reported above was conducted in training environments, and for the most part the participants were officers. The high validity and reliability that associate evaluation has demonstrated make the technique potentially valuable as an added measure for use in making personnel decisions in combat and combat support units. However, the situational constraints in training differ from those in line units, and the applicability and acceptability of peer evaluation in the two settings could lead to very different conclusions about the use of such evaluations.

OBJECTIVES

The present research was designed to assess the acceptability of a peer evaluation technique among enlisted men in combat and combat support units. Specifically, the objectives were to (1) measure soldier attitudes toward peer evaluations using the Unit Associate Rating Questionnaire (UARQ); (2) gather information regarding the relative acceptability of various evaluation techniques for use in making decisions as to promotion or separation from the service; and (3) determine whether non-training units pose any feasibility problems for the administration of associate ratings.

2 → p. 10
-1-

METHOD

SAMPLE

Soldiers participating in the research were members of the 27th Brigade, 42d Division of the New York State National Guard. The research was conducted at Fort Drum, N.Y. during the 27th Brigade's two-week annual training period. However, unlike the training units in previous work on peer evaluations (Officer Basic Training Course, Army Individual Training, and ROTC), participating units were regularly constituted National Guard units whose personnel were assigned on a long-term basis.

The five platoons, one from each of five different companies of the 27th Brigade, were representative of various combat and combat support units. Table 1 shows the companies and number of individuals for each platoon. Because the research procedures were carried out in the field during breaks in the schedule, situational factors made it impossible for all the designated soldiers to participate in the procedures, and resulted in smaller Ns than were anticipated. A total of 79 enlisted men were included.

Table 1
PLATOONS SAMPLED AND
NUMBER FROM EACH

Platoon	N
Co. D, 102nd Medical Battalion	19
Co. B, 1/127th Armor Battalion	11
Headquarters and Headquarters Co., 1/108th Infantry	20
Co. A, 1/174th Infantry	17
Combat Support Co., 2/108th Infantry	<u>12</u>
	79

VARIABLES

Data were collected using three separate instruments. Soldiers first filled out a biographic survey form designed to obtain demographic information on each participant's educational, vocational, and military background. The form also contained two items which asked soldiers to state their opinion of the acceptability of various assessment methods used in the military for decisions regarding (1) promotions and (2) separation. Associate rating was included as one of the assessment techniques. The various techniques were then rank ordered according to how desirable they were thought to be for career decisions.

The Guardsmen next participated in an associate evaluation procedure. As used here, associate evaluations are modified full nominations where the raters are required to choose from among the other members of their platoon the six soldiers whom they consider best at getting the job done and the six soldiers whom they consider to be the worst at getting the job done. These choices were recorded on optical scanning answer sheets.

Following the associate evaluations procedure, participants were required to complete the UARQ, a modified and expanded version of the Associate Rating Questionnaire (ARQ) which has been used in previous associate rating research on officers (Downey, 1975). The UARQ incorporated questions from the ARQ that also applied to enlisted men; it eliminated items dealing specifically with areas peculiar to officers and trainees. The UARQ also contained an extensive number of items related to the National Guard climate. The present analysis focused on opinions about the validity and use of peer rating techniques. Specifically, respondents were asked about their attitudes regarding the ability of associates to produce valid ratings, what the rating scores should be used for, how rating scores might affect their day-to-day performance, and how long rating scores should be kept on an individual's record.

ANALYSIS

A two-way analysis of variance with repeated measures was performed to determine if differential unit membership had an effect on responses to the UARQ. Associate evaluations were correlated with attitude measures on the UARQ to determine if any relationship existed between ratings and their acceptability. Another correlational analysis was performed between peer evaluations and self-evaluations. Finally, t tests were calculated comparing the mean response for each item against a score which represented a neutral opinion. Significant results for any of the t tests would indicate a general opinion which was either positive or negative toward the item. Also presented were opinions about various methods of making promotion or separation decisions.

RESULTS AND DISCUSSION

The associate ratings from each of the five platoons were found to be highly reliable, and within the range reported by other researchers (Hollander, 1957). The adjusted split-half reliability coefficients ranged from a low of .80 for the medical platoon to .93 for the infantry platoon.

Associate rating scores were correlated with each of the first 21 UARQ items, except items 17 and 18 which were self-ratings. Associate rating scores were also correlated with the rank given to the associate evaluation technique in comparison with other methods of making promotion or separation decisions. None of the coefficients was statistically significant, indicating that there was no relationship between attitudes toward peer ratings and the actual associate evaluation score received. The fact that attitudes and ratings were independent indicates that soldiers answering the questionnaire or ranking the various personnel decision methods must not have taken into account how they felt their own evaluations would turn out. Respondents evidently had some idea of what their own associate ratings would be; correlation between associate ratings and self-ratings was $r = .41$ ($p < .01$). Correlation was $r = .46$ ($p < .01$) between associate ratings and how respondents estimated others would rate them.

Tables 2 and 3 present the proportion of respondents indicating whether various personnel decision methods are acceptable or unacceptable for promotion or separation decisions, and the ranking of each method for each decision. Associate evaluations were considered acceptable for use in promotion decisions; however, opinion about use of the ratings in separation decisions was divided. For both promotion and separation decisions, associate evaluation was ranked in the middle of the distribution of methods. With the exception of supervisors' ratings, which ranked third for both types of decision, the evaluation methods that were ranked above associate rating have an objective quality. "Evaluation of your work" ranked first and "Skill Qualification Test" ranked second for both decisions; soldiers seem to prefer evaluation based on performance. The high ranking of "Supervisor's rating" may be the result of the feeling that supervisors are in the best position to evaluate an individual's work. Other items in the UARQ dealt with attitudes toward supervisors and are not relevant to the present analysis.

Five different units had been sampled under the assumption that various combat and combat support units might view the acceptability of associate ratings differently. Table 4 presents the results of a two-way analysis of variance to determine if a difference between units existed in responses to the UARQ items 1 through 20. The results of the analysis indicate that while participants responded to the items differentially, there was no significance difference between responses of the five platoons, nor was the interaction between items and units significant.

Table 2

RANKING BY SOLDIERS FROM BEST (1) TO WORST (11)
 METHOD FOR DETERMINING PROMOTIONS; AND
 PERCENTAGE OF SOLDIERS INDICATING
 ACCEPTABILITY OF EACH METHOD

Rank	Item	Acceptable:	
		"Yes"	"No"
1	Evaluation of one's work	89%	11%
2	Skill Qualification Test	83	17
3	Supervisor's rating	82	18
4	Completion of professional development training	81	19
5	Rating by one's associates	73	27
6	MOS test	65	35
7	Time in grade	60	40
8	Board decision	39	61
9	Commander's decision	39	61
10	Getting assigned slot which has higher rank associated with it	33	67
11	First sergeant's choice	29	71

Table 3

RANKING BY SOLDIERS FROM BEST (1) TO WORST (9) METHOD
FOR DETERMINING INVOLUNTARY SEPARATION; AND PERCENTAGE
OF SOLDIERS INDICATING ACCEPTABILITY OF EACH METHOD

Rank	Item	Acceptable:	
		"Yes"	"No"
1	Evaluation of one's work	89%	11%
2	Skill Qualification Test	78	22
3	Supervisor's rating	73	27
4	Rating by one's associates	55	45
5	MOS test	58	42
6	Time in grade	39	61
7	Board decision	55	45
8	Commander's decision	47	53
9	First sergeant's choice	31	69

Table 4

RESULTS OF TWO-WAY ANALYSIS OF VARIANCE FOR UARQ
ITEMS 1 - 20 AND PARTICIPATING PLATOONS

Source	Sum of Squares	df	MS	F
<u>Between Subjects</u>				
Platoons	8.00	3	2.67	.88 ns
Subjects within Platoons	195.81	65	3.01	
<u>Within Subjects</u>				
Items	142.53	19	7.50	5.27 **
Items x Platoons	75.22	57	1.32	.93 ns
Subjects x Platoons x Items	1757.77	1235	1.42	

** $p < .001$

Table 5 shows the means and standard deviations for responses to items 1 through 21 of the UARQ, excluding items 17 and 18 which were handled separately. Also indicated is whether responses deviated significantly from undecided (a score of 3). The Appendix presents the proportion of respondents who chose each response for items 1-21. The overall pattern of results shows a positive attitude toward the use of associate ratings as a part of soldiers' records. The mean values for items 1 and 2 indicate that participants in this study felt that associate ratings could be of value in predicting future work performance and that the situations upon which judgments were made in this case were adequate. The very positive response to these items is contradictory to the findings of Mohr (1975) in Officer Basic Course settings where raters were asked to evaluate leadership potential. Officers there felt that associate ratings were not predictive of performance, and that the situations upon which a rating would be based were not sufficient for valid and reliable ratings.

Items 3, 4, and 5 dealt with the effects that associate evaluations might have on performance or on interpersonal relations within the work unit. Replies indicate that raters were undecided about the effects of peer ratings. Respondents indicated a slight disagreement with the idea that the associate evaluation procedure would affect interpersonal relations or cause an individual to work harder.

Items 6, 7, 8, 9, and 10 addressed various aspects of the premise that a person can judge which members of a group of associates are performing well and which are performing poorly. Responses to item 6 were relatively positive, indicating that although one may work closely with only a few individuals one can still judge the effectiveness of most members of a platoon. (Content of item 6 has been abbreviated in Table 5.) Responses to question 7 indicated agreement with the idea that associate ratings should be based on long-term association. Responses to item 8 were also positive with regard to individuals being capable of determining who was not performing well. Responses to item 9, which asked if peers rate friends high regardless of ability, were about evenly divided and indicated that in many cases respondents were being objective in their evaluations. Item 10 showed some disagreement with the statement that not many people could judge the performance of others, a result consistent with the over-all assessment that individuals are capable of making rating decisions, given a long enough period of time working together.

Questions 11, 12, 13, and 14 proposed various uses for associate ratings. Respondents seemed to be undecided about what specific uses the ratings should be put to, although positive about including ratings as part of the total record. This result is consistent with the finding reported earlier that peer evaluation was acceptable but was ranked in the middle of a list of evaluation methods. Item 15 asked respondents to choose how long an associate rating should remain in their records, the continuum being from not wanting associate ratings to be in the record at all to keeping them there until replaced by a subsequent rating.

Table 5

MEANS AND STANDARD DEVIATIONS AND t -TEST RESULTS
FOR ITEMS 1 - 21 OF THE UARQ

Item ^a	Mean ^b	SD	r^c	t^d
1. Ratings Predict Future Performance	2.61	.98	.03	3.51 ***
2. Length of Time Adequate	2.54	1.10	.12	3.70 ***
3. Ratings Produce Self Change	2.76	1.17	-.19	1.83
4. Ratings Could Cause Defensive Behavior	3.08	1.34	.04	.50
5. Ratings Cause Harder Work	3.19	1.32	-.09	1.29
6. One Can Judge Others	2.79	1.29	.01	1.47
7. Ratings Should Be Based on Long Experience	2.63	1.43	-.16	2.29 *
8. Peers Know Who Isn't Performing	2.53	1.03	.09	3.62 ***
9. Peers Rate Friends High	2.81	1.18	-.08	1.43
10. Ratings Are Not Accurate	3.08	1.11	-.08	.61
11. Ratings Should Be Used To Select For Training	2.70	1.23	.21	2.12 *
12. Ratings Should Be Used for Promotions	2.90	1.22	.11	.74
13. Ratings Should Be Part Of The Total Record	2.57	1.21	.03	3.11 ***
14. Ratings Should Be Used In Separation Decisions	3.01	1.32	.02	.09
15. Ratings Should Be Kept For A Time ^e	3.17	1.49	-.10	.99
16. How Many In Platoon Do You Work With? ^f	2.26	1.03	.14	6.29 ***
19. Ratings Affect Morale	2.52	1.35	-.10	3.09 ***
20. Ratings Affect Cohesiveness	2.72	1.34	.09	1.93
21. Ratings Affect Work Output	2.27	1.23	-.03	5.25 ***

^a Items 17 and 18 excluded; these were self-ratings.

^b Lower values indicate greater agreement; a value of 3.00 indicates an average neutral position.

^c Positive correlations between peer ratings and UARQ items indicate that individuals with higher peer scores tend to have greater agreement with the UARQ item.

^d t compares mean of UARQ with neutral position.

^e Lower values indicate less time or "do not want on the record."

^f Lower values indicate larger number of close co-workers.

* $p < .05$

** $p < .01$

*** $p < .001$

The mean response was between keeping them till promoted and keeping them indefinitely, with decreasing weight over time. The reply to this item also contrasts with Mohr's (1975) results in which officer basic students overwhelmingly preferred not to have the scores in their records.

The final items of the peer evaluation section of the UARQ asked what effect periodic associate ratings, if used in career progression decisions, would have on morale, cohesiveness, and work output. Responses were positive with respect to both morale and work output but indicated that ratings would have no effect on cohesiveness. Responses to later UARQ items indicated that cohesiveness was already quite high.

The results of the present research are inconsistent with those found by Mohr (1975) using officer basic students, in which strong negative feelings were expressed toward associate evaluation as an acceptable technique. An explanation for this marked difference in results lies in the nature of the settings in which the studies took place. An Officer Basic Course brings together a diverse group of individuals who are learning how to adjust to a new life style. During the OBC period, officers are expected to learn many of the skills that they will be called upon to use during the rest of their military careers. The introduction of the associate evaluation procedure into this adjustment and learning period was evidently felt to be inappropriate by the OBC students who participated in the evaluations and also, as Mohr reports, by the school personnel, who treated the evaluation procedure as an added chore. In the National Guard, evaluation is expected by the trainees and is carried out by first sergeants and commanders. It is therefore not unexpected that Guardsmen favored having their associates, who are in many cases their friends, rate them.

The difference in impact of the ratings between officers in training and enlisted men might also explain why Downey (1975) found that when he manipulated instructions about associate evaluation, those officers who responded with positive attitudes on the ARQ tended to have received high rating scores from their associates. The Officer Basic Course students who felt they could benefit were most probably more willing to express positive opinions about associate ratings. In the present research, although associate ratings were accepted, there was no relationship between attitudes and rating scores. Guardsmen are possibly less threatened by evaluation, and therefore can more freely express their opinions, independent of the outcome of the evaluation.

CONCLUSIONS

Generally, the National Guardsmen who participated in the present research demonstrated willingness to accept associate evaluation as a part of their records. This positive response was inconsistent with negative attitudes shown toward peer evaluation by officers in training situations.

Most of the Guardsmen felt that associate evaluations could predict future performance and should be made part of individual records. They also felt that they had enough experience with the men they were rating to be able to make valid decisions. Guardsmen were undecided about specific uses for ratings, especially if the ratings were to be used for separation. Some opinions were expressed that use of rating scores in making career progression decisions would have a positive effect on morale and work output.

A problem that occurred repeatedly and that would probably occur in most non-training situations was the impossibility of coordinating schedules so that all those who were to participate in the data collection were actually present. The present study was conducted during the annual training period and no platoon sampled was able to produce all the subjects requested because of other demands.

The positive outcome of the present research indicates that the National Guard would accept the use of peer evaluation in career decisions. This result should not be generalized too far. The nature of the present sample (average age 25, education post-high school, length of service over two years) and the nature of individual involvement in the Guard would preclude inferring similar attitudes in a Regular Army sample. These results do warrant investigation of the acceptability of associate ratings in active Army combat and combat support units. While demographic characteristics would differ markedly from those of the National Guard, group structures would be similar enough to hypothesize acceptance in Regular Army settings.

SUMMARY

National Guardsmen (N = 79) participated in an associate rating procedure in which they were asked to rate those in their platoon who were best and worst at getting the job done. They then completed the Unit Associate Rating Questionnaire (UARQ) which sampled attitudes about the value and acceptability of associate evaluation procedures, and also completed a biographic survey. Guardsmen expressed generally favorable attitudes about associate evaluation, in contrast to previous findings in officer populations. In their opinion, associate ratings were based on enough and sufficiently varied experiences with other members of the platoon to be valuable for predicting future performance. While the Guardsmen were willing to accept associate ratings as part of their records, they were neutral with respect to specific uses of the rating scores in career decisions. There was no significant difference between attitudes of the various combat and combat support units in the sample. The use of associate evaluation as a career progression tool in non-training units was not precluded by the present findings, and merits further evaluative investigation.

REFERENCES

- Bolton, W. L. An application of constant sum paired comparison techniques to peer ratings at an Army ROTC summer camp. Unpublished master's thesis, University of Tennessee (Knoxville), 1971.
- Downey, R. G. Associate evaluations: Nominations vs ratings. ARI Technical Paper 253. September 1974. (AD 785 550)
- Downey, R. G. Associate evaluations: Improving field acceptance. ARI Research Memorandum 75-5. July 1975.
- Downey, R. G. Associate nominations in the U.S. Army officer training environment: The Ranger course. ARI Research Problem Review 76-8. October 1976.
- Downey, R. G., and Duffy, P. J. Various methodological approaches to peer evaluations. Paper presented at the 21st Conference on the Design of Experiments in Army Research, Development and Testing, Washington, D.C., October 1975.
- Downey, R. G., Medland, F. F., and Yates, L. G. Evaluation of a peer rating system for predicting subsequent promotion of senior military officers. ARI Research Memorandum 76-7. April 1976.
- Hollander, E. P. The reliability of peer nominations under various conditions of administration. Journal of Applied Psychology, 1957, 41, 85-90.
- Mohr, E. S. Acceptability of associate ratings at branch basic schools. ARI Technical Paper 268. October 1975. (AD A017 437).

APPENDIX

	Page
Proportion of respondents choosing each response alternative	15

Table A-1

PROPORTION OF RESPONDENTS CHOOSING EACH RESPONSE
ALTERNATIVE ITEMS 1-14 UARQ

	Completely Agree	Mostly Agree	Undecided	Mostly Disagree	Completely Disagree
1. Ratings Predict Future Performance	8.9	46.8	24.0	16.5	3.8
2. Length Of Time Adequate	12.7	49.4	16.5	15.2	6.3
3. Ratings Produce Self Change	11.4	36.7	30.4	8.9	12.7
4. Ratings Could Cause Defensive Behavior	11.4	30.4	16.5	22.8	19.0
5. Ratings Cause Harder Work	8.9	30.4	16.5	22.8	21.5
6. One Can Judge Others	12.7	43.0	11.4	19.0	13.9
7. Ratings Should Be Based On Long Experience	29.1	27.8	6.3	25.3	11.4
8. Peers Know Who Isn't Performing	12.7	43.0	27.8	11.4	5.1
9. Peers Rate Friends High	10.1	39.2	19.0	21.5	10.1
10. Ratings Are Not Accurate	7.6	26.6	22.8	35.4	7.6
11. Ratings Should Be Used To Select For Training	16.5	36.7	19.0	17.7	10.1
12. Ratings Should Be Used For Promotions	11.4	32.9	24.0	19.0	12.7
13. Ratings Should Be Part Of The Total Record	17.9	42.3	15.4	15.4	9.0
14. Ratings Should be Used In Separation Decisions	12.8	28.2	24.4	15.4	19.2

Table A-2

PROPORTION OF RESPONDENTS CHOOSING
EACH RESPONSE ALTERNATIVE
ITEMS 15-16 UARQ

Item 15:	Shouldn't Be On Record	One Year	Until Promoted	Indefi- nitely	Until Replaced
IF RATINGS ARE A PART OF YOUR RECORD, HOW LONG DO FAVOR ITS USE?	20.2	15.2	16.5	21.5	26.6
Item 16:	Entire Platoon	Squad	3 or 4 Others	One Other	Work Alone
HOW MANY MEMBERS OF YOUR PLATOON DO YOU WORK CLOSELY WITH?	28.2	30.8	33.3	3.8	3.8

Table A-3
PROPORTION OF RESPONDENTS CHOOSING EACH RESPONSE
ALTERNATIVE ITEMS 17-18 UARQ

	Upper 1/5	Mid- upper 1/5	Mid 1/5	Mid- lower 1/5	Lower 1/5
	1	2	3	4	5
17. WHERE DO YOU RATE YOUR WORK COM- PARED TO OTHERS IN YOUR PLATOON?	35.9	34.6	21.8	5.1	2.6
18. WHERE DO YOU THINK YOUR SCORE ACTUALLY FALLS?	16.7	39.7	33.3	5.1	5.1

Table A-4
PROPORTION OF RESPONDENTS CHOOSING EACH RESPONSE
ALTERNATIVE ITEMS 19-21 UARQ

	Very Positive Effect	2	No Effect	4	Very Negative Effect
19. RATINGS AFFECT MORALE	29.1	22.8	27.8	6.3	13.9
20. RATINGS AFFECT COHESIVENESS	24.0	19.0	31.6	11.4	13.9
21. RATINGS AFFECT WORK OUTPUT	34.2	25.3	27.8	3.8	8.9